Soo Wan Kim
MACS 30000
Short Paper 2: Observational Study Proposal

1. Introduction

In the United States, Thanksgiving and Christmas are traditionally considered holidays of goodwill and giving. Part of the broader holiday season, they are associated with gift shopping, charity drives, and spending time with family and friends. While the holidays' association with online sales are well-known, few studies have focused on other manifestations of "holiday spirit" online. Popular media sharing websites such as YouTube contain rich and accessible stores of information on the public's media consumption patterns over time. They present a yet unseized opportunity to measure differences in individual users' preoccupations and interests – even concern for the less fortunate – as a function of proximity to certain holidays. This proposed study will use YouTube news channel viewing data to address the following question: **Do American news viewers display greater interest in international humanitarian crises, as measured by volume of views and likes on related videos on YouTube, during periods closer to Thanksgiving and Christmas?**

The findings of this study will have implications not only for social scientists seeking to understand media consumption behavior or holidays as a driver of broad social trends, but also for media companies seeking to optimize their programming and for nonprofits seeking to target their fundraising outreach to maximize donations yield.

2. Data

The data cover video news segments uploaded on the twelve most viewed American news channels on YouTube. Namely, they are the official channels for CNN, USA TODAY, HLN, NBC News, ABC News, Associated Press, Fox News, CBS News, the New York Times, the Huffington Post, PBS NewsHour, and MSNBC (Social Blade; VidStatsX). Each channel has over 100,000 total views, and together they have a total of 405,042 unique uploads since 2005. A rough estimate using the latest hundred CNN videos suggests that 2% of the total, or approximately 8,100 videos relate to international humanitarian crises specifically. However, these recent videos are heavily centered on the US election, and so the actual total may be higher.

These channels are ideal for several other reasons. First, they represent mainstream news agencies widely familiar to the American public and they cover a broad range of domestic and international topics without specializing in any one topic or locality. As such, they appeal to a much broader base of American news consumers than would local news stations or more specialized services such as Bloomberg or the Wall Street Journal. Second, they represent some of the most competitive and highly regarded news services in the United States. Consequently, the relative quality of the videos can be assumed to be largely consistent within and across channels over time. Third, at least within the past few years each channel has been uploading new videos continuously, typically every day. Thus, considering that all the channels are at least five years old and most are ten to eleven years old, this data represents several uninterrupted yearly cycles of news production and viewership.

3. Methods and Feasibility

Definition of international humanitarian crises

For this project, I will define international humanitarian crises as events or ongoing series of events that 1) happened or are happening outside of the US, 2) primarily affected non-US citizens, 3) caused 100 or more casualties, and 4) are possible for the average American to impact in a way that is concrete and measurable, and using only simple methods, such as monetary donations to relief agencies.

Following the fourth condition, the crises should be viewed as at least mitigatable if not resolvable in the short term, with the outcomes being largely dependent on the immediate efforts of individuals, not on complex long-term policy commitments by governments. News coverage on tragic events that are too difficult for ordinary citizens to meaningfully impact may simply frustrate the viewer, leading to desensitization and lack of interest over time. While the meaning of "meaningfully impact" is subjective, I will argue that refugee crises, most natural disasters, and certain health crises fit the above definition, as ordinary citizens can safely donate funds to help rescue and deliver necessities to individual victims with reasonable chance of success. On the other hand, terrorist attacks, military conflicts, and human rights violations outside of the US do not fit the definition, as in these cases most non-state actors are powerless to intervene without taking drastic measures such as taking up arms or illegally entering territory held by a hostile regime.

Data collection

YouTube has a flexible and well-documented API, which makes it easy to scrape targeted data from videos while adhering to YouTube's terms of service. Registering for a Google Developers account will allow access to the API, which will then allow me request the publication date, the number of views, and the number of likes from each video in a list of videos at a single point in time.

The main challenge will be to isolate the relevant videos. To do this, I will search within channels via the API using a list of key terms. The list will be generated by taking random samples of videos across channels, isolating relevant videos within samples, then identifying terms that appear frequently or necessarily in the titles (e.g. "Syria," "hurricane," "casualties"). After compiling unique videos via the search terms, I will look through the entire set individually to weed out irrelevant videos or refine the search as needed. This method, though crude, will not take more than a few weeks to complete.

Analysis

Each observation in the final dataset will include the following characteristics related to a single video: total number of likes, total number of views, number of days since video publication until the next Thanksgiving, number of days since video publication until the next Christmas, number of days since video publication since the last Thanksgiving, and the number of days since video publication since the last Christmas. Because views accumulate over time, I will also include time-weighted versions of the number of views and likes per video. To select the weighting method, I will use the Statistics tool under the video player options to study viewing patterns over time. Also, because views and likes may not be normally distributed, I may use log transformations or bin the data depending on the initial summary statistics. Finally, I

will conduct a simple regression analysis using R or Stata to test for causal relationships between proximity to the two holidays and the absolute or relative numbers of views and likes.

4. Caveats

First, the dataset will not perfectly capture views and likes as a function of time of the year. As mentioned, videos continue to accumulate views long after they are first published. However, an initial glance at the trends using the Statistics tool suggests that views largely peak in the days immediately following publication and subside at an accelerating rate over time, such that the total number of views are still representative of relative popularity. I will study the YouTube API documentation in greater depth to look for ways to measure views and likes at different time points rather than in a "snapshot."

Second, although this study seeks to measure *American* interest in international humanitarian crises, there is nothing to prevent the millions of non-American YouTube users around the world from watching American news channels on the site. There will probably be a significant number of foreign views and likes per video. In addition, an unknown proportion of American viewers will prefer foreign news channels or prefer them equally to domestic channels. Thus, there is a non-trivial issue of construct validity. However, because foreign channels do not cover US domestic news as extensively as American channels, it is likely that American viewers will tend to prefer domestic channels over foreign for their daily news.

Third, YouTube users are not representative of the general American public. Survey data indicates that 57% of YouTube users are male; 38% are younger than 30, and 30% are between 30 and 49. At the same time, an estimated 10% of American adults – over 20 million individuals – get news from YouTube (Gottfried & Shearer 2016).

Fourth, there may be some channel-specific effects. For example, CNN and Fox have long had popular cable networks in addition to printed media, whereas the Huffington Post and the New York Times are still mostly known for their newspapers. These characteristics may affect video content and viewers' usage of the channels. In addition, differences in viewers' ideological preferences may lead to different viewing trends and uploading trends across channels. To test, I will analyze the data at both an aggregate level and at the channel level and adjust the methods and/or interpretation as needed.

Fifth, there may be temporal effects. For example, channels may have been less active and less popular initially after opening. Also, viewing and uploading patterns may have been significantly affected by the domestic economic and political climate in certain periods, particularly during the financial crisis and during election years. However, because the data will cover a period of over seven years, these effects should be balanced out to some extent. Lastly, the news agencies themselves might choose to release more humanitarian crisis videos during certain parts of the year based on their own expectations. To account for this, I will measure the trends for humanitarian crisis video uploads across channels and adjust the methods and/or interpretation of the data as necessary.

References

Gottfried, Jeffrey, and Elisha Shearer. "News Use Across Social Media Platforms 2016." Pew Research Center. Accessed October 23, 2016. http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/.

"Top 500 YouTubers News Channels." *Social Blade*. Accessed October 23, 2016.
     http://socialblade.com/youtube/top/category/news/mostviewed.

"YouTube Top 100 Most Viewed News & Politics Video Producers." *VidStatsX*. Accessed
     October 23, 2016. https://vidstatsx.com/youtube-top-100-most-viewed-news-politics#.